

Model klasyfikacji, podział LOSOWY

Accuracy 0.6, F1 0.33, Precision 1, Recall 0.2, Specificity 1

Ocena wyników modelu

Accuracy 60%: niska ogólna trafność modelu (40% wszystkich przypadków jest błędnie klasyfikowanych).

F1 Score 0.33: słaby balans między Precision a Recall, głównie przez niski Recall.

Precision 100%: model nie generuje fałszywych pozytywów (FP = 0).

Każde przewidziane "tak" jest pewne.

Recall 20%: model wykrywa 20% rzeczywiście pozytywnych przypadków.

Przewidziane "nie" nie są pewne.

Specificity 100%: Model poprawnie rozpoznaje wszystkie rzeczywiste „nie” (TN) i nigdy nie myli „nie” z „tak” (FP = 0).

Wniosek

Niska ogólna trafność modelu.

Model jest skrajnie ostrożny i unika False Positives (Specificity = 100%), ale może jednocześnie popełniać wiele False Negatives (niedostateczny Recall) i pomijać większość rzeczywistych pozytywnych przypadków, co czyni go mało użytecznym, mimo idealnej precyzji.

Model klasyfikacji, podział STRATYFIKOWANY

Accuracy 0.8, F1 0.33, Precision 1, Recall 0.2, Specificity 1

Ocena wyników modelu

Accuracy 80%: dobra ogólna trafność modelu (tylko 20% wszystkich przypadków jest błędnie klasyfikowanych).

F1 Score 0.33: słaby balans między Precision a Recall, głównie przez niski Recall.

Precision 100%: model nie generuje fałszywych pozytywów (FP = 0).
Każde przewidziane "tak" jest pewne.

Recall 20%: model wykrywa 20% rzeczywiście pozytywnych przypadków.
Przewidziane "nie" nie są pewne.

Specificity 100%: Model poprawnie rozpoznaje wszystkie rzeczywiste „nie” (TN) i nigdy nie myli „nie” z „tak” (FP = 0).

Wniosek

Dobra ogólna trafność modelu.

Model jest skrajnie ostrożny i unika False Positives (Specificity = 100%), ale może jednocześnie popełniać wiele False Negatives (niedostateczny Recall) i pomijać większość rzeczywistych pozytywnych przypadków, co czyni go mało użytecznym, mimo idealnej precyzji.

Model STRATYFIKOWANY jest lepszy niż LOSOWY, ponieważ przy zachowaniu pełnej Precyzji i Specyficzności, zapewnia wyraźnie wyższą ogólną trafność klasyfikacji tj. Accuracy 80%. Oba modele mają niski Recall (model wykrywa tylko 1 na 5 rzeczywiście pozytywnych przypadków), co ogranicza ich wartość praktyczną, ale model STRATYFIKOWANY jest zdecydowanie stabilniejszy i bardziej godny zaufania.